APPENDIX J

HISTORICAL FREQUENCY OF OCCURRENCE ASSESSMENT OF PIPES THROUGH AN EMBANKMENT

J-1. Introduction. A reliability analysis model can be developed using historic data collected at a specific project or similar projects. The historic data would consist of known unsatisfactory performance events and the time of occurrence of the events. The historic frequency of occurrence model can be developed using the Weibull, Exponential, or Poisson distributions or a hazard function model. Examples of developing a historic frequency of occurrence model using the Weibull and Exponential distributions for pipes through an embankment are presented as follows.

J-2. Weibull Distribution.

a. The Weibull distribution is the most general method for developing a historic frequency of occurrence model and is used predominantly in industrial engineering. This method of reliability analysis is used to determine the probability of failure, reliability, or hazard function of a component with time using unsatisfactory performance data. A probability density function is fitted to the failure data versus time. This method of reliability engineering is based solely on unsatisfactory performance data without using any knowledge of the engineering properties or condition of the component. This method of reliability analysis is presented in the textbook "Reliability in Engineering Design" by K. C. Kapur and L. R. Lamberson.

b. The probability density function f(t) for the Weibull distribution is:

$$f(t) = \frac{b}{\alpha}\left[\frac{t}{\alpha}\right]^{b-1} \exp\left[-\left[\frac{t}{\alpha}\right]^{b}\right] \tag{1}$$

where

b is the shape parameter.

α is the characteristic life.

t is time.

F(t) is the cumulative distribution function, the probability that the system will fail by the time t or the probability of failure. F(t) is given as follows:

$$F(t) = 1 - \exp\left[-\left[\frac{t}{\alpha}\right]^{b}\right] \tag{2}$$

R(t) is the reliability function, the probability that the system will not fail by time t.

$$R(t) = \exp\left[-\left[\frac{t}{\alpha}\right]^{b}\right] \tag{3}$$

As can be seen from an examination of Equations 2 and 3, the reliability and probability of failure are related by the following equation:

$$R(t) = 1 - F(t) \tag{4}$$

The hazard function h(t) is the rate of failure at time t. The hazard function is the probability of failure in a given time period, usually one year, given that that item has not already failed. Simply stated, if there are ten items and one item fails in year one and another item fails in year two, the hazard function for year two is the number of items that failed in year two divided by the number of items existing at the beginning of year two. For this simple example the hazard function for year two would be:

$$h(t) = h(2) = \frac{1}{10 - 1} = 0.11$$

The hazard function for the Weibull distribution is given as

$$h(t) = \frac{b}{\alpha}\left[\frac{t}{\alpha}\right]^{b-1} \tag{5}$$

The Weibull distribution has the following characteristics: For a shape parameter of b = 1, the Weibull distribution becomes the exponential distribution, which gives a constant hazard function with an equal rate of failure in any year. For b = 2, the Weibull distribution becomes the Rayleigh distribution, which gives a linearly increasing hazard function. For b < 1, the hazard function decreases with time, giving a decreasing rate of failure with time. For b > 1, the hazard function increases with time, giving an increasing rate of failure with time. A b value of 1 would be representative of the occurrence of a random event, such as scour occurring adjacent to a structure, erosion, an earthquake, or an accident. Deterioration of sheetpiling could be represented by a b value between 1 and 2. For any Weibull distribution, there is a 63.2 percent probability that failure will occur before the characteristic life and a 37.8 percent probability that failure will occur after the characteristic life. Put another way, 63.2 percent of the components will fail by the characteristic life and 37.8 percent will not fail.

c. The term “failure” leads one to think in terms of a catastrophic event such as the failure of a large dam resulting in grave consequences. However, when addressing failures of components of a dam, failure could mean something as insignificant as an electric motor not starting when needed. So, to avoid the confusion that could result from the use of the term “failure”, the term “unsatisfactory performance” is used.

(1) From past known unsatisfactory performance events a table can be compiled of the unsatisfactory performance events that have occurred and the time of occurrence of that event. Each unsatisfactory performance event needs to be listed in the table according to the time of occurrence. The table needs to include the date the item was put into operation, the date of the unsatisfactory event, the time in years to the unsatisfactory event, and the corrected percentage of the components that have experienced unsatisfactory performance. The corrected percentage of components that have experienced unsatisfactory performance is a median rank correction to the data. The median rank correction adjusts the data such that the probability that the unsatisfactory performance event occurs before time t is 50%. The median rank correction is made using the following equation:

$$Median\ Rank\ Correction\,(\%) = \left[\frac{j-0.3}{N+0.4}\right]*100 \qquad (6)$$

where

j is the number of the item that had unsatisfactory performance, i.e., first item, second item, ……jth item

N is the sample size that is, the total of the items having unsatisfactory performance plus the items having satisfactory performance. To use the Weibull distribution to develop a historical frequency of occurrence model, the number of items that have had unsatisfactory performance and those that did not have unsatisfactory performance must be known.

(2) The data from the table (time and corrected percentage of components that have experienced unsatisfactory performance) is plotted on Weibull graph paper. A straight line is passed through the data points to determine b and α, where b is the slope of the straight line and α is the time at which 63.2 percent of the items have experienced unsatisfactory performance. The straight line is located by sight. While a least-squares fit or some other numerical method could be used, the sight method should be used based on the following quote from "Reliability in Engineering Design" by Kapur and Lamberson: "A least-square fitting procedure could be used; however, this somewhat defeats the purpose, namely, the ease of graphical estimation." The sight method of fitting the straight line to the data requires the data to be plotted and allows engineering judgment to be applied to the straight line fit. A mathematic method of fitting the straight line to the data, while being mathematically more accurate, can lead to serious errors, especially if the data is not plotted.

(3) An error in the historic database would be that it does not include all unsatisfactory performance events due to inadequate records. Thus, the historic database will tend to underestimate the actual number of events that have occurred and thus those that will occur in the future.

J-3. Weibull Distribution Example. In a particular levee system, there are 23 large diameter corrugated metal pipes (CMP) going through the levee. These pipes are connected to gravity drain structures. Three of these CMP, have failed during flood events causing water to flow into the protected area. Two CMP, failed during the 1973 flood and one failed during the 1993 flood. Refer to the Table J-1 to see when the pipes were installed. Using the Weibull distribution, calculate the hazard function in the year 2003 given that there is a flood. Calculate the hazard function for the years 2013 and 2023 given that there is a flood in those years.

Table J-1. Selected CMP Installation and Failure Information

	Year Installed	**Year of Failure**	**Time to Failure**	**Corrected % failure**
1st Failure	1950	1973	23	3.0
2nd Failure	1950	1973	23	7.3
3rd Failure	1950	1993	43	11.5

Solution:

First, use Equation (6) to find the corrected percent failure of the pipes:

For 1st pipe failure in 1973: Corrected % failure $= \dfrac{1-0.3}{23+0.4} \times 100 = 3$

For 2nd pipe failure in 1973: Corrected % failure $= \dfrac{2-0.3}{23+0.4} \times 100 = 7.3$

For 3rd pipe failure in 1993: Corrected % failure $= \dfrac{3-0.3}{23+0.4} \times 100 = 11.5$

On the Weibull graph paper, plot on the x-axis the time to failure. On the y-axis, plot corrected percent failure. Plot the points: (23,3), (23,7.3) and (43,11.5) on the Weibull graph paper (See Figure J-1).

Draw a best-fit line through the data points. The slope of the best-fit line is b = 1.4 (the slope b is obtained by matching the slope of the best fit line to the Weibull Slope indicator in the upper left hand corner of the Weibull distribution plot). The best fit line crosses 63.2% at 230 years so $\alpha = 230$

Calculate the hazard function for year 2003 using Equation (5): for t = 53 (2003 – 1950)

$$h(t) = \frac{1.4}{230}\left[\frac{53}{230}\right]^{(1.4-1)} = 0.0034$$

For the year 2013, t = 2013 – 1950 = 63 years. Using Equation (5):

$$h(t) = \frac{1.4}{230}\left[\frac{63}{230}\right]^{(1.4-1)} = 0.0036$$

For the year 2023, t = 73 years:

$$h(t) = \frac{1.4}{230}\left[\frac{73}{230}\right]^{(1.4-1)} = 0.0038$$

Over the twenty-year time period from 2003 to 2023 the hazard function increased from 0.0034 to 0.0038. This is the hazard function for a single corrugated metal pipe. Since there are 23 of the corrugated metal pipes in the levee system, a system hazard function must be developed. Guidance on calculating system reliability is given in ETL 1110-2-547.

J-4. Exponential Distribution. The Exponential distribution is a continuous probability density function. The probability density function (pdf) for this distribution is:

$$f(t) = \lambda e^{-\lambda t} \tag{7}$$

where

λ is the expected number of events in a time period

t is the time period

The exponential distribution can be used to model any process that is considered to be a random event such as floods, earthquakes, or scour. An important application for the exponential distribution is to model the time to the first occurrence of a Poisson event and the time between Poisson events. Poisson events are events that occur randomly in time or space and have an equally likely probability of occurrence in any unit time or space increment Δt. A Poisson distribution is used when the probability of a certain number of events in a given time period is desired.

In order to determine the probability that an event has occurred by time t, or within the interval (0,t) the cumulative distribution function (cdf) is used. The cdf for the exponential distribution is obtained by integrating the probability density function:

$$F(t) = \int_o^t \lambda e^{-\lambda t} = 1 - e^{-\lambda t} \tag{8}$$

Equation (8) provides the probability that the first event (or failure) occurs by time t. So F(t) is the probability of failure. To determine the probability that a failure will occur in the interval t_1 to t_2, the following equation can be used:

$$\Pr(\text{failure in interval } t_1 \text{ to } t_2) = F(t_2) - F(t_1) \tag{9}$$

The probability that no failure (or event) occurs by time t is represented by the reliability function R(t). R(t) is the complement of the cumulative distributive distribution and has the following equation:

$$R(t) = 1 - F(t) = e^{-\lambda t} \qquad (10)$$

The hazard function h(t) for the exponential distribution can be obtained by dividing the probability density function by the reliability function. The hazard function is the conditional probability of failure in the next unit time increment given that the system has survived to time t. The hazard function is given as the following equation:

$$h(t) = \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda \qquad (11)$$

Since h (t) = λ, it can be seen that the exponential distribution has a constant hazard rate. Historical data can be used to determine λ. The information required to calculate λ is a data set of items that have been in service with and without failures. Of the items that failed, it is necessary to know the times to failure and times in service. λ can be calculated using the following equation:

$$\lambda = \frac{F}{T} \qquad (12)$$

where

F = number of failures; and

T = the total time in service for failed and unfailed items.

J-5. Exponential Example. In a particular levee system, there are 23 large corrugated metal pipe (CMP) gravity drains. Three of these CMPs have failed during a flood. Two CMP failed during the 1973 flood and one failed during the 1993 flood. Refer to Table J-2 to see when all of the CMPs were installed. Using the exponential distribution, calculate the hazard function.

Solution:

The total time in service (T) for all of the pipes is $\sum_{i=1}^{23} (\textit{Life} \text{ of Pipe}) = 1086$ pipe-years

Use Equation (12) to calculate λ:

$$\lambda = \frac{3\,\textit{failures}}{1086\,\textit{pipe years}} = 0.0028$$

The hazard function for a single pipe is calculated as

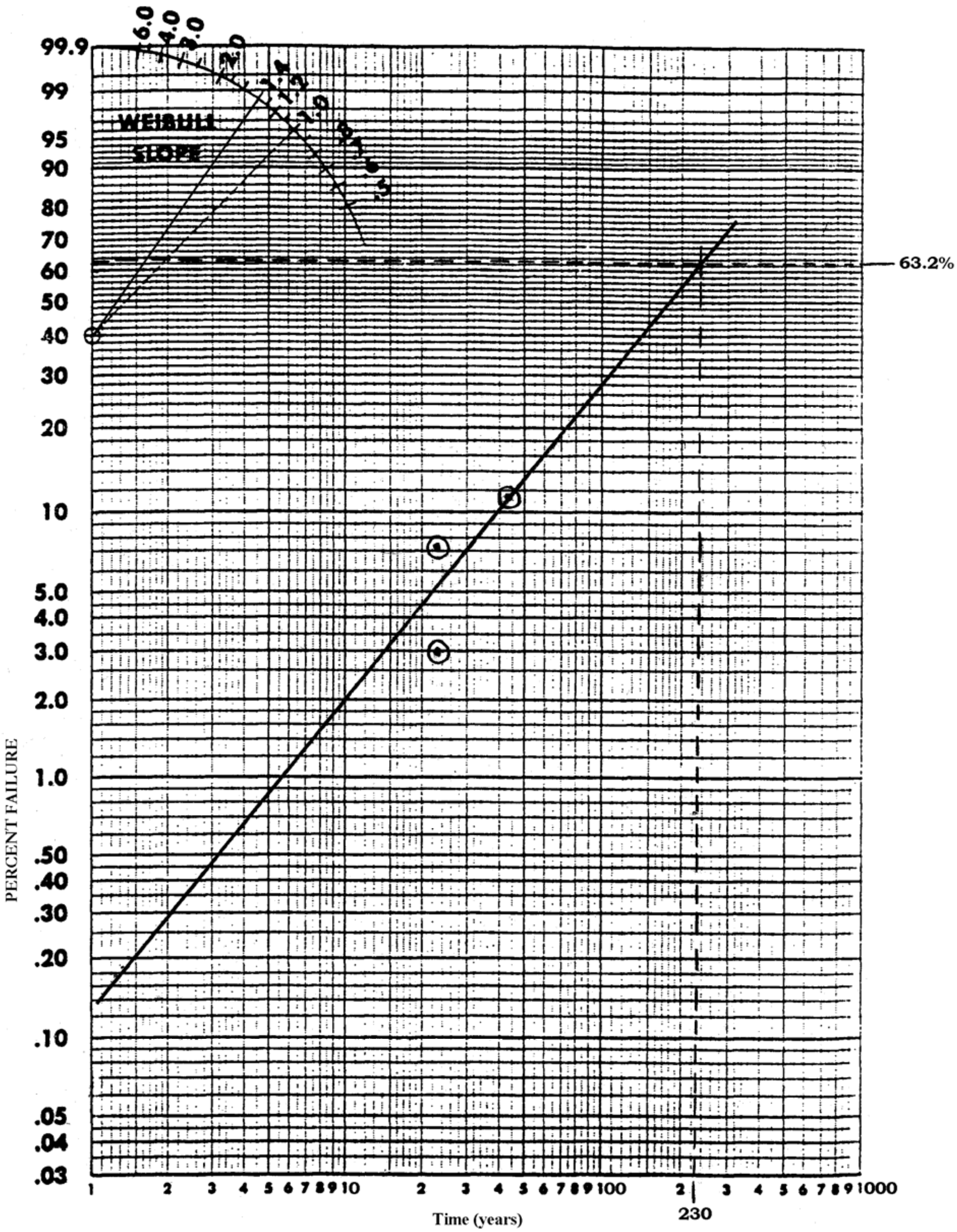
$h(t) = \lambda = 0.0028$

Using the Exponential distribution, the hazard rate is 0.0028 and is constant with time.

J-6. Comparison Between the Weibull Distribution and the Exponential Distribution. The Weibull distribution example and the exponential distribution example are the same problem, yet they yield different hazard rates. The difference between the two examples is that the Weibull example has an increasing hazard function while the exponential example has a constant hazard function. The Weibull distribution's hazard rate increases from 0.0034 in year 2003 to 0.0036 in the year 2013. The exponential example has a constant hazard rate for every year of 0.0028. If the b value of the Weibull example had equaled 1, the Weibull example and the exponential example would have yielded the same result. When b = 1 for a Weibull distribution, the Weibull distribution becomes the Exponential distribution. A b value of 1 would represent the occurrence of a random event. Since pipe failure is dependent on deterioration of the pipe (which leads to failure) it is legitimate to assume that the value of b would be between 1 and 2. If pipe failure were only dependent on a flood, it would be reasonable to assume that pipe failure follows an exponential distribution.

Table J-2. Total CMP Installation and Failure Information

Pipe #	Date Installed	Date Failed	Life of Pipe (to year 2003 or failure)
1	1950	1973	23
2	1950	1973	23
3	1948		55
4	1950		53
5	1950		53
6	1950		53
7	1950		53
8	1950		53
9	1950		53
10	1950		53
11	1960		43
12	1952		51
13	1952		51
14	1950	1993	43
15	1953		50
16	1956		47
17	1956		47
18	1956		47
19	1956		47
20	1956		47
21	1956		47
22	1956		47
23	1956		47
Sumation (Total time in service,T)			1086



Weibull Distribution Plot
Figure J-1